AD-A208 084

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution is unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: Naval Ocean Systems Center; 6b. OFFICE SYMBOL (If applicable): NOSC	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State and ZIP Code): San Diego, CA 92152-5000	7b. ADDRESS (City, State and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b. OFFICE SYMBOL (If applicable): ONR	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code): Math Info Sciences Division, Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	AGENCY ACCESSION NO.
0601153N	ST95	RR01405	DN305 106

11. TITLE (Include Security Classification)

AUTOREGRESSIVE SPECTRAL ESTIMATION IN ADDITIVE NOISE

12. PERSONAL AUTHOR(S)

D. F. Gingras, E. Masry

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Professional paper	FROM TO	March 1989	

16. SUPPLEMENTARY NOTATION

17. COSATI CODES		
FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

bearing estimation
array processing
statistics
time series analysis

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The estimation of the spectral density of a discrete-time stationary Gaussian autoregressive (AR) process from a finite set of noisy observations is considered. A modified spectral estimator based on the high-order Yule-Walker equations is considered. Joint asymptotic normality of this spectral estimator is established; a precise asymptotic expression for the covariance matrix of the limiting distribution is obtained. The special case of AR(1) plus noise is considered in some detail.

Published in *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Volume 36, No. 4, April 1988.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: [X] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT [ ] DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE PERSON: D. F. Gingras	22b. TELEPHONE (Include Area Code): (619) 553-2050; 22c. OFFICE SYMBOL: Code 733

DD FORM 1473, 84 JAN

83 APR EDITION MAY BE USED UNTIL EXHAUSTED
ALL OTHER EDITIONS ARE OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

19. ABSTRACT (Continued)

T-ASSP/36/4//19015

Autoregressive Spectral Estimation in Additive Noise

Donald F. Gingras
Elias Masry

Reprinted from
IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING
Vol. 36, No. 4, April 1988

Autoregressive Spectral Estimation in Additive Noise

DONALD F. GINGRAS, MEMBER, IEEE, AND ELIAS MASRY, FELLOW, IEEE

Abstract—The estimation of the spectral density of a discrete-time stationary Gaussian autoregressive (AR) process from a finite set of noisy observations is considered. A modified spectral estimator based on the high-order Yule-Walker equations is considered. Joint asymptotic normality of this spectral estimator is established; a precise asymptotic expression for the covariance matrix of the limiting distribution is obtained. The special case of AR(1) plus noise is considered in some detail.

I. Introduction

In this paper, the problem of estimating the spectral density of a discrete-time stationary Gaussian autoregressive (AR) process from a finite set of noisy observations is considered. AR spectral estimation techniques are widely used in signal processing applications. They are meaningful when the underlying process is well modeled by an AR random process and the resulting spectral estimate will exhibit, for small sample size, increased resolution (see Kaveh and Cooper [1] and Beamish and Priestly [2]) over that obtained by the nonparametric smoothed periodogram. In practice, the observation process has a noise component and the usual noise-free AR spectral estimate is no longer adequate, especially when the signal-to-noise ratio is low. A modified estimate is considered and its asymptotic statistical properties are developed.

Let $X = \{X_n\}_{n=-\infty}^{\infty}$ be a real-valued autoregressive process of order p, AR(p), with spectral density $\phi(\lambda)$ and let $\{W_n\}_{n=-\infty}^{\infty}$ be a white noise process. The observed process $Y = \{Y_n\}_{n=-\infty}^{\infty}$ is defined by $\{Y_n = X_n + W_n\}$ and let $\psi(\lambda)$ denote its spectral density. The problem being considered is the estimation of the spectral density $\phi(\lambda)$ of the AR(p) process X from a finite set of noisy observations $\{Y_n\}_{n=1}^{N}$.

For the noise-free case, the asymptotic statistics for the estimates of the AR parameters were rigorously established by Mann and Wald [3]. They proved that the estimates were consistent, that their limit distribution was normal and calculated the asymptotic covariance. Akaike [4], starting with the asymptotic results of Mann and Wald [3] for the parameters' estimates, and assuming that the order p is known, proved that the asymptotic distribution for the corresponding AR spectral estimate is normal and calculated its covariance. Berk [5], under the assumption that the order of the AR process goes to infinity as the number of observations tends to infinity, proved that the distribution of the AR spectral estimate is asymptotically normal with asymptotic variance identical to that of a truncated periodogram.

The problem of AR parameter estimation for the AR plus noise case was first examined by Walker [6] who evaluated the asymptotic efficiency and variance for the parameter estimate of a first-order AR process. Pagano [7], noting that the correct model for an AR(p) plus noise process is an autoregressive-moving average (ARMA) process of order (p, p), developed strongly consistent and efficient AR parameter estimates through the use of nonlinear regression. Lee [8] examined the multivariate AR plus noise case and proved asymptotic normality for estimates of the AR parameters based on the high-order Yule–Walker equations. However, his results, which depend on a central limit theorem of Billingsley [9], are not substantiated as shown in Section V. For the univariate AR(p) plus noise case, asymptotic normality for estimates of the AR parameters, based on the high-order Yule–Walker equations, was established by Gingras [10].

In this paper, we consider the estimation of the spectral density $\phi(\lambda)$ of an autoregressive process AR(p) from a finite set of noisy observations; the spectral estimate $\hat{\phi}_N(\lambda)$ utilizes AR parameters' estimates, based on the high-order Yule–Walker equations. The goal of the paper is to establish the asymptotic statistics of the spectral estimate $\hat{\phi}_N(\lambda)$. The organization of the paper is as follows. In Section II, the basic assumptions are set down; the estimates for the parameters and the spectral density $\hat{\phi}_N(\lambda)$ are defined. The main results are presented in Section III where the joint asymptotic normality of $\hat{\phi}_N(\lambda)$ is established along with an expression for the asymptotic covariance. In Section IV, the special case of AR(1) plus noise is considered in some detail, and a performance comparison to the noise-free case is presented. Appendix A contains the proofs of certain lemmas used in the derivation of the main results, and Appendix B contains detailed calculations of variance and covariance expressions. Throughout this paper, vectors are denoted by lower case letters and matrices are denoted by upper case letters; both are indicated by boldface type.

II. Preliminaries

Throughout this paper, the signal $X = \{X_n\}_{n=-\infty}^{\infty}$ is assumed to be a stationary Gaussian AR(p) process. Such

Manuscript received September 16, 1986; revised September 28, 1987. The work of D. F. Gingras was supported by the Office of Naval Research, Statistical Signal Analysis Program. The work of E. Masry was supported by the Office of Naval Research under Contract N00014-84-K-0042.

D. F. Gingras is with the Naval Ocean Systems Center, Code 733, San Diego, CA 92152.

E. Masry is with the Department of Electrical Engineering and Computer Sciences, University of California at San Diego, La Jolla, CA 92023.

IEEE Log Number 8719015.

a process has the representation

$$X_n - a_1 X_{n-1} - \cdots - a_p X_{n-p} = \epsilon_n; \quad n = 0, \pm 1, \cdots \tag{1}$$

where the innovations process $\{\epsilon_n\}$ is a sequence of independent identically distributed Gaussian random variables $\mathfrak{N}(0, \sigma_\epsilon^2)$. The AR parameters $\{a_j\}_{j=1}^p$ are constrained such that the polynomial

$$A(z) = -\sum_{j=0}^{p} a_j z^j \; (a_0 = -1)$$

has no zeros inside the closed unit circle $\{z: |z| \leq 1\}$. The spectral density $\phi(\lambda)$ of the process X is given by

$$\phi(\lambda) = \frac{\sigma_\epsilon^2/2\pi}{\left|1 - \sum_{j=1}^{p} a_j e^{-ij\lambda}\right|^2}. \tag{2}$$

The noise process $W = \{W_n\}_{n=-\infty}^{\infty}$ is assumed to be a sequence of independent identically distributed Gaussian random variables $\mathfrak{N}(0, \sigma_W^2)$. Moreover, the processes X and W are assumed to be independent.

The observation process $Y = \{Y_n\}_{n=-\infty}^{\infty}$ is defined by

$$Y_n = X_n + W_n; \quad n = 0, \pm 1, \cdots.$$

The spectral density $\psi(\lambda)$ of the process $\{Y_n\}$ is given by

$$\psi(\lambda) = \sigma_W^2/2\pi + \phi(\lambda). \tag{3}$$

The process Y is then an ARMA(p, p) series with representation (see Walker [6] and Pagano [7])

$$\begin{aligned} Y_n - a_1 Y_{n-1} - \cdots - a_p Y_{n-p} \\ = \epsilon_n + W_n - a_1 W_{n-1} - \cdots - a_p W_{n-p} \end{aligned} \tag{4}$$

for all n. Denote the covariance sequence of the series Y by $\{r_k\}$ where $r_k = E[Y_n Y_{n-k}]$. Multiplying (4) through by Y_{n-k} and taking expectations, we obtain the Yule-Walker (Y-W) equations:

$$r_0 - a_1 r_1 - \cdots - a_p r_p = \sigma_\epsilon^2 + \sigma_W^2 \tag{5}$$

$$r_k - a_1 r_{k-1} - \cdots - a_p r_{k-p} = -a_k \sigma_W^2; \quad \text{for } 1 \leq k \leq p \tag{6}$$

$$r_k - a_1 r_{k-1} - \cdots - a_p r_{k-p} = 0; \quad \text{for } k \geq p+1. \tag{7}$$

The set of p equations in (7) corresponding to $k = p + 1, \cdots, 2p$ is often referred to as the high-order Yule-Walker equations. We express this set in matrix form as

$$\boldsymbol{R}\boldsymbol{a} = \boldsymbol{b} \tag{8}$$

where the $(p \times p)$ covariance matrix $\boldsymbol{R}$ is defined by

$$\boldsymbol{R} \equiv \begin{bmatrix} r_p & r_{p-1} & \cdots & r_1 \\ r_{p+1} & r_p & \cdots & r_2 \\ \vdots & \vdots & & \vdots \\ r_{2p-1} & r_{2p-2} & \cdots & r_p \end{bmatrix} \tag{9}$$

and the $(p \times 1)$ vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are defined by

$$\boldsymbol{a} = (a_1, a_2, \cdots, a_p)^T$$
$$\boldsymbol{b} = (r_{p+1}, r_{p+2}, \cdots, r_{2p})^T.$$

Given a finite set of noisy observations $\{Y_n\}_{n=1}^N$, $N > 2p$, we estimate the covariance sequence $\{r_k\}$ using

$$\hat{r}_{N,k} = \begin{cases} (1/N) \sum_{n=1}^{N-|k|} Y_n Y_{n+|k|}, & |k| \leq N-1 \\ 0, & |k| > N-1. \end{cases} \tag{10}$$

When the elements of the matrix $\boldsymbol{R}$ and vector $\boldsymbol{b}$ are replaced by their corresponding estimates (10), the estimated matrix and vector will be denoted by $\hat{\boldsymbol{R}}_N$ and $\hat{\boldsymbol{b}}_N$, respectively. Using the high-order Yule-Walker equations (8), we define the vector estimate $\hat{\boldsymbol{a}}_N$ of the AR parameters as the solution of the equation

$$\hat{\boldsymbol{R}}_N \hat{\boldsymbol{a}}_N = \hat{\boldsymbol{b}}_N. \tag{11}$$

To estimate the AR spectral density $\phi(\lambda)$, we require estimates of the AR parameters such as those formed by (11) and an estimate of σ_ϵ^2. An estimate of σ_ϵ^2 can be obtained by first using (5) to provide an estimate of $\sigma_\epsilon^2 + \sigma_W^2$ and then using one of the equations (6), say $k = p$, to obtain an estimate for σ_W^2. This yields

$$\hat{\sigma}_{N,\epsilon}^2 = -\sum_{j=0}^{p} \hat{a}_{N,j} \hat{r}_{N,j} - (1/\hat{a}_{N,p}) \sum_{j=0}^{p} \hat{a}_{N,j} \hat{r}_{N,p-j} \tag{12}$$

where $\hat{a}_{N,0} = -1$ since $a_0 = -1$. Note that for large N, $\hat{a}_{N,p}$ is necessarily nonzero since $\hat{a}_{N,p}$ converges almost surely to a_p (see Lemma 1 in Section III below), and $a_p \neq 0$ since the order of the process X is p. The spectral estimate $\hat{\phi}_N(\lambda)$ is now defined by

$$\hat{\phi}_N(\lambda) = \frac{\hat{\sigma}_{N,\epsilon}^2/2\pi}{\left|1 - \sum_{j=1}^{p} \hat{a}_{N,j} e^{-ij\lambda}\right|^2}. \tag{13}$$

For the sake of compactness of notation, we let

$$\boldsymbol{\theta} = (\sigma_\epsilon^2, a_1, \cdots, a_p)^T \tag{14a}$$

be the vector of parameters to be estimated and $\hat{\boldsymbol{\theta}}_N$ be its estimate

$$\hat{\boldsymbol{\theta}}_N = (\hat{\sigma}_{N,\epsilon}^2, \hat{a}_{N,1}, \cdots, \hat{a}_{N,p})^T. \tag{14b}$$

In the subsequent development of asymptotic statistical properties for the parameter and spectral density estimates, we make use of the following vectors and matrices, some of which have been defined previously, but are presented here for convenient reference:

$$c = (r_0, r_1, \cdots, r_{2p})^T \tag{15a}$$

$$b = (r_{p+1}, r_{p+2}, \cdots, r_{2p})^T \tag{15b}$$

$$a = (a_1, a_2, \cdots, a_p)^T \tag{15c}$$

$$\tilde{a} = (a_p, a_{p-1}, \cdots, a_1)^T \tag{15d}$$

$$d = (r_0, r_1, \cdots, r_p)^T \tag{15e}$$

$$\tilde{d} = (r_p, r_{p-1}, \cdots, r_0)^T \tag{15f}$$

$$U(\lambda) = [u_{k,j}(\lambda)]_{k,j=0}^{2p} \tag{15g}$$

$$u_{k,j}(\lambda) = e^{i(k+j)\lambda} + e^{i(k-j)\lambda}$$

$$\mathbf{0} = (0, 0, \cdots, 0)^T \tag{15h}$$

where the dimension of the null vector $\mathbf{0}$ will be clear by the context in which it is used.

In the next section, we make repeated use of the following results, regarding convergence in probability $\xrightarrow{P}$ and convergence in distribution $\xrightarrow{\mathcal{L}}$, the proofs of which can be found in Rao [11].

Proposition 1: For each integer $N \geq 1$, let η_N and ζ_N be random vectors and let α and β be constant vectors; then

$$\eta_N \xrightarrow{\mathcal{L}} \eta, \; \zeta_N \xrightarrow{P} \mathbf{0} \Rightarrow \eta_N^T \zeta_N \xrightarrow{P} 0 \tag{16a}$$

$$\eta_N \xrightarrow{P} \alpha, \; \zeta_N \xrightarrow{P} \beta \Rightarrow \eta_N^T \zeta_N \xrightarrow{P} \alpha^T \beta \tag{16b}$$

$$|\eta_N - \zeta_N| \xrightarrow{P} \mathbf{0}, \; \zeta_N \xrightarrow{\mathcal{L}} \zeta \Rightarrow \eta_N \xrightarrow{\mathcal{L}} \zeta. \tag{16c}$$

III. Statistical Properties

In this section, we establish the joint asymptotic normality of the spectral estimate $\hat{\phi}_N(\lambda)$ and provide a closed-form expression for its covariance matrix (Theorem 3). Since the statistical properties of the covariance estimates $\hat{r}_{N,k}$ (10) are well known (see Brillinger [12]), the method of analysis adopted in this paper is to establish asymptotic equivalence in probability between the error vector $(\hat{a}_N - a)$ and appropriate linear combinations of $\{\hat{r}_{N,k} - r_k\}_{k=0}^{2p}$ (Lemma 2); similarly, we establish such an equivalence relationship for the innovation variance's error $\hat{\sigma}_{N,\epsilon}^2 - \sigma_\epsilon^2$ (Lemma 3). Using these relationships, the statistical properties of the errors $(\hat{a}_N - a)$ and $\hat{\sigma}_{N,\epsilon}^2 - \sigma_\epsilon^2$ can be established from those of $\{\hat{r}_{N,k}\}$. These two lemmas and their derivations are the crux of the analysis leading to the main result (Theorem 3). For the sake of readability, we have delegated the complex derivation of Lemma 3 (as well as that of Lemma 2) to Appendix A.

A. Statistics of Parameters' Estimates

It should be noted that in [10], the asymptotic normality of the AR parameters' estimates $(\hat{a}_N)$ was established. However, this result does not lead directly to the *joint* asymptotic normality of $\hat{\sigma}_{N,\epsilon}^2$ and $\hat{a}_N$ in view of the dependence of $\hat{\sigma}_{N,\epsilon}^2$ on a_N as well as on the covariance estimate $\{\hat{r}_{N,j}\}$. Lemmas 2 and 3, which were not developed in [10], provide an appropriate asymptotic reduction of $\hat{\theta}_N = (\hat{\sigma}_{N,\epsilon}^2, \hat{a}_N^T)^T$ in terms of the covariance estimates $\{\hat{r}_{N,j}\}$. This reduction makes it possible to establish the asymptotic normality of $\hat{\theta}_N$ and provide a convenient expression for its asymptotic covariance Σ. In principle, one could consider expressing $\hat{\sigma}_{N,\epsilon}$ solely in terms of $\hat{a}_N$ (by (12); this would require writing the $\{\hat{r}_{N,j}\}$ in terms of $\hat{a}_N$) and then applying the results of [10] for $\hat{a}_N$ to establish asymptotic normality of $\hat{\theta}_N$; this approach does not appear to be simpler in terms of the complexity of the derivations. In this subsection, we establish the asymptotic distribution for $\hat{\theta}_N$. First, we present the asymptotic distribution of the covariance estimates of (10) as established by Brillinger [12, p. 256].

Theorem 1: For the observation process Y, the covariance estimate $\hat{r}_{N,k}$ converges in quadratic mean to r_k as $N \to \infty$ with covariance

$$\lim_{N\to\infty} N \operatorname{cov}\{\hat{r}_{N,k}, \hat{r}_{N,j}\} = 2\pi \int_{-\pi}^{\pi} \{e^{i\lambda(j+k)} + e^{i\lambda(j-k)}\} \psi^2(\lambda)\, d\lambda. \tag{17}$$

Moreover, the standardized estimates $N^{1/2}(\hat{r}_{N,1} - r_1)$, $N^{1/2}(\hat{r}_{N,2} - r_2), \cdots, N^{1/2}(\hat{r}_{N,m} - r_m)$ are asymptotically jointly normal with zero means and asymptotic covariance given by the right side of (17).

Thus, the standardized vector estimate $N^{1/2}(\hat{c}_N - c)$ of (15a) is asymptotically multivariate normal with zero mean and covariance matrix given by

$$\lim_{N\to\infty} \operatorname{cov}\{N^{1/2}(\hat{c}_N - c), N^{1/2}(\hat{c}_N - c)^T\} = 2\pi \int_{-\pi}^{\pi} U(\lambda)\, \psi^2(\lambda)\, d\lambda. \tag{18}$$

The following lemma, whose proof is delegated to Appendix A, establishes the almost sure convergence of $\hat{R}_N^{-1}$ and $\hat{a}_{N,k}$ as $N \to \infty$. We remark that R^{-1} exists (see Gersch [13]).

Lemma 1: $\hat{R}_N^{-1}$ exists almost surely for large N and converges to R^{-1} as $N \to \infty$ almost surely. Moreover, the AR parameter estimate $\hat{a}_{N,k}$ converges almost surely to a_k as $N \to \infty$ for $k = 1, \cdots, p$.

We next construct a random vector $N^{1/2} z_N$ that is equivalent in probability to the high-order Y-W AR parameter estimate vector $N^{1/2}(\hat{a}_N - a)$. The proof is given in Appendix A.

Lemma 2: Define the vector random variable z_N by

$$z_N = R^{-1} D(\hat{c}_N - c) \tag{19}$$

and the matrix D by

$$D \equiv \begin{bmatrix} 0 & -a_p & -a_{p-1} & \cdots & -a_1 & 1 & 0 & & \cdots & 0 \\ 0 & 0 & -a_p & \cdots & & -a_1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & & & & & & & & \vdots \\ 0 & 0 & \cdots & 0 & -a_p & & \cdots & & -a_1 & 1 \end{bmatrix}.$$

Then

$$N^{1/2}\left|(\hat{a}_N - a) - z_N\right| \xrightarrow{P} \mathbf{0} \text{ as } N \to \infty. \tag{20}$$

We previously established an estimate (12) for the variance σ_ϵ^2. The next lemma, whose proof is given in Appendix A, establishes the existence of a random variable $N^{1/2}\zeta_N$ that is equivalent in probability to $N^{1/2}(\hat{\sigma}_{N,\epsilon}^2 - \sigma_\epsilon^2)$.

Lemma 3: Define the random variable ζ_N by

$$\zeta_N = h^T(\hat{c}_N - c) \tag{21}$$

where

$$h = \left(-(-1, a^T, 0^T) - (1/a_p)(\bar{a}^T, -1, 0^T)\right.$$
$$+ \left(d^T + (1/a_p)\,\bar{d}^T\right)\begin{bmatrix} 0^T \\ R^{-1} \end{bmatrix} D$$
$$\left. + (1/a_p)^2((-1, a^T)\,\bar{d})\,[R^{-1}D]_p\right)^T \tag{22}$$

and $[A]_p$ denotes the pth row of the matrix A. Then

$$N^{1/2}\left|(\hat{\sigma}_{N,\epsilon}^2 - \sigma_\epsilon^2) - \zeta_N\right| \xrightarrow{P} 0 \text{ as } N \to \infty. \tag{23}$$

The following result establishes the asymptotic normality of the vector estimate $\hat{\theta}_N$. Its proof relies on the equivalence relationships of Lemmas 2 and 3.

Theorem 2: The standardized $((p + 1) \times 1)$ vector estimate $N^{1/2}(\hat{\theta}_N - \theta)$ is asymptotically multivariate normal, that is,

$$N^{1/2}(\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} \mathfrak{N}(0, \Sigma) \text{ as } N \to \infty$$

where

$$\Sigma \equiv \begin{bmatrix} v^2 & s^T \\ s & G \end{bmatrix} \tag{24}$$

and

$$v^2 \equiv 2\pi \int_{-\pi}^{\pi} h^T U(\lambda)\, h \psi^2(\lambda)\, d\lambda \tag{25}$$

$$s \equiv 2\pi \int_{-\pi}^{\pi} R^{-1} D U(\lambda)\, h \psi^2(\lambda)\, d\lambda \tag{26}$$

$$G \equiv 2\pi \int_{-\pi}^{\pi} R^{-1} D U(\lambda)\, D^T (R^{-1})^T \psi^2(\lambda)\, d\lambda. \tag{27}$$

Note that the dimensions of the matrices Σ and G are $((p + 1) \times (p + 1))$ and $(p \times p)$, respectively.

Proof: By definition, we have

$$N^{1/2}(\hat{\theta}_N - \theta) = N^{1/2}\begin{bmatrix} \hat{\sigma}_{N,\epsilon}^2 & -\sigma_\epsilon^2 \\ \hat{a}_N & -a \end{bmatrix};$$

by Lemma 2, we have

$$N^{1/2}\left|(\hat{a}_N - a) - z_N\right| \xrightarrow{P} \mathbf{0} \text{ as } N \to \infty \tag{28}$$

where z_N is defined in (19). Also, by Lemma 3, we have

$$N^{1/2}\left|(\hat{\sigma}_{N,\epsilon}^2 - \sigma_\epsilon^2) - \zeta_N\right| \xrightarrow{P} 0 \text{ as } N \to \infty \tag{29}$$

where ζ_N is defined in (21). Thus, it follows from (28) and (29) that

$$N^{1/2}\left|(\hat{\theta}_N - \theta) - \begin{bmatrix} h^T \\ R^{-1}D \end{bmatrix}(\hat{c}_N - c)\right| \xrightarrow{P} \mathbf{0} \text{ as } N \to \infty.$$

From Theorem 1, we have that $N^{1/2}(\hat{c}_N - c)$ is asymptotically multivariate normal; therefore, by (16c), it follows that $N^{1/2}(\hat{\theta}_N - \theta)$ converges in distribution to a multivariate normal vector with zero mean and covariance matrix Σ which, by (18), is given by

$$\Sigma = 2\pi \int_{-\pi}^{\pi} \begin{bmatrix} h^T \\ R^{-1}D \end{bmatrix} U(\lambda)\left[h, D^T (R^{-1})^T\right] \psi^2(\lambda)\, d\lambda. \tag{30}$$

The result follows. □

Theorem 2 establishes convergence in distribution of the vector estimate $N^{1/2}(\hat{\theta}_N - \theta)$. It should be noted that Σ is the covariance matrix of the asymptotic distribution. This does not necessarily imply that the covariance of the vector estimate $N^{1/2}(\hat{\theta}_N - \theta)$ converges to Σ of (30) (see Serfling [14, p. 20]). For the sake of simplicity, however, we refer to Σ as the asymptotic covariance matrix instead of the more accurate expression "covariance matrix of the asymptotic distribution." Similar terminology is used in connection with the asymptotic properties of the spectral estimate.

The components of the asymptotic covariance matrix Σ can be evaluated using expression (3) for the spectral density $\psi(\lambda)$. The detailed evaluation of these terms is presented in Appendix B.

B. Statistics of Spectral Density Estimate

We now establish the joint limiting distribution of the spectral estimates $\hat{\phi}_N(\lambda_1), \cdots, \hat{\phi}_N(\lambda_r)$ at r distinct frequencies $\lambda_1, \cdots, \lambda_r$. We make use of the following theorem on nonlinear transformations of asymptotically normal variates (Serfling [14, Theorem 3.3A]) which we recast in our notation in the next lemma.

Lemma 4: Let $\hat{\theta}_N = (\hat{\theta}_{N,1}, \cdots, \hat{\theta}_{N,m})^T$ be jointly asymptotically normal random variables with mean vector θ and covariance matrix Σ, i.e.,

$$N^{1/2}(\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} \mathfrak{N}(0, \Sigma) \text{ as } N \to \infty.$$

Let $g(x) = (g_1(x), \cdots, g_r(x))^T$, $x = (x_1, \cdots, x_m)^T$ be a vector-valued function for which each component $g_k(x)$ is real valued and has a nonzero differential $g_k(\theta; \tau)$ at $x = \theta$:

$$g_k(\theta; \tau) = \sum_{j=1}^{m} \left.\frac{\partial g_k}{\partial x_j}\right|_{x=\theta} \tau_j$$

for any vector $\tau = (\tau_1, \cdots, \tau_m)^T$.

Put

$$\Gamma = \left[\left.\frac{\partial g_k}{\partial x_j}\right|_{x=\theta}\right]_{\substack{k=0,\cdots,r \\ j=0,\cdots,m}}$$

Then

$$g(\hat{\theta}_N) = (g_1(\hat{\theta}_N), \cdots, g_r(\hat{\theta}_N))^T$$

is jointly asymptotically normal with mean $g(\theta)$ and covariance matrix $\Gamma\Sigma\Gamma^T$, i.e.,

$$N^{1/2}(g(\hat{\theta}_N) - g(\theta)) \xrightarrow{\mathcal{L}} \mathfrak{N}(0, \Gamma\Sigma\Gamma^T) \text{ as } N \to \infty.$$

Theorem 3: The standardized spectral density estimates $N^{1/2}\hat{\phi}_N(\lambda_1), \cdots, N^{1/2}\hat{\phi}_N(\lambda_r)$ are jointly asymptotically normal, as $N \to \infty$, with means $\phi(\lambda_1), \cdots, \phi(\lambda_r)$ and covariance matrix $\Gamma\Sigma\Gamma^T$, i.e.,

$$N^{1/2}[\hat{\phi}_N(\lambda_1) - \phi(\lambda_1)], \cdots,$$
$$N^{1/2}[\hat{\phi}_N(\lambda_r) - \phi(\lambda_r)] \xrightarrow{\mathcal{L}} \mathfrak{N}(0, \Gamma\Sigma\Gamma^T) \quad (31)$$

where the $r \times (p+1)$ matrix Γ is given by

$$\Gamma = (\rho(\lambda_1), \cdots, \rho(\lambda_r))^T \quad (32a)$$

and the $(p+1) \times 1$ column vector $\rho(\lambda)$ is defined by

$$\rho(\lambda) = \phi(\lambda)[1/\sigma_\epsilon^2, e^T(\lambda)]^T \quad (32b)$$

with

$$e(\lambda) = 2 \operatorname{Re}\left[\frac{e^{i\lambda}}{A(e^{i\lambda})}, \frac{e^{i2\lambda}}{A(e^{i\lambda})}, \cdots, \frac{e^{ip\lambda}}{A(e^{i\lambda})}\right]^T. \quad (32c)$$

Proof: Let $m = p + 1$. By Theorem 2, we have that

$$N^{1/2}(\hat{\theta}_N - \theta) \xrightarrow{\mathcal{L}} \mathfrak{N}(0, \Sigma) \text{ as } N \to \infty \quad (33)$$

where θ and $\hat{\theta}_N$ are given by (14). Now the spectral density $\phi(\lambda_k)$ is a real-valued function of the parameter vector μ, i.e., $\phi(\lambda_k) = \phi_k(\mu)$ where

$$\phi_k(\mu) = \frac{\mu_0/2\pi}{\left|1 - \sum_{j=1}^{p} \mu_j e^{ij\lambda_k}\right|^2}.$$

It is seen that the partial derivatives

$$\frac{\partial\phi_k(\mu)}{\partial\mu_j}, \qquad j = 0, 1, \cdots, p$$

exist and are continuous in a neighborhood of $\mu = \theta$. Moreover, the gradient vector

$$\rho_k(\theta) = \left[\frac{\partial\phi_k(\theta)}{\partial\theta_0}, \cdots, \frac{\partial\phi_k(\theta)}{\partial\theta_p}\right]^T$$

is given by (32b), i.e., $\rho_k(\theta) = \rho(\lambda_k)$. We note that $e(\lambda)$ of (32c) actually depends on θ via $A(e^{i\lambda})$ which is a function of $\{a_j\}_{j=1}^p$ [cf. (14a)]. Thus, $\rho(\lambda)$ of (32b) depends implicitly on θ. Since at least the first component of $\rho_k(\theta)$ is clearly nonzero, it follows that the differential

$$g_k(\theta; \tau) \triangleq \rho_k^T(\theta)\tau$$

is nonzero for each $k = 1, \cdots, r$. Lemma 4 now applies with the identification $g_k(x) = \phi_k(x)$, and the result follows from (33). □

Theorem 3 establishes the joint asymptotic normality of the scaled spectral estimate errors $N^{1/2}[\hat{\phi}_N(\lambda) - \phi(\lambda)]$ and provides explicit expressions for the variance/covariance of the joint asymptotic distribution. Using the decomposition (24) for Σ, we have that the asymptotic covariance function of $N^{1/2}[\hat{\phi}_N(\lambda) - \phi(\lambda)]$ and $N^{1/2}[\hat{\phi}_N(\nu) - \phi(\nu)]$ appearing in the asymptotic distribution (31) is given by

$$\begin{aligned}\operatorname{acov}(\lambda, \nu) = \phi(\lambda)\phi(\nu)\Big\{&(v^2/\sigma_\epsilon^4) \\ &+ e^T(\lambda)s/\sigma_\epsilon^2 + s^Te(\nu)/\sigma_\epsilon^2 \\ &+ e^T(\lambda)Ge(\nu)\Big\}, \end{aligned} \quad (34a)$$

and if we set $\lambda = \nu$, we obtain the asymptotic variance avar (λ) of $N^{1/2}[\hat{\phi}_N(\lambda) - \phi(\lambda)]$ as

$$\begin{aligned}\operatorname{avar}(\lambda) = \phi^2(\lambda)\Big\{&(v^2/\sigma_\epsilon^4) + 2e^T(\lambda)s/\sigma_\epsilon^2 \\ &+ e^T(\lambda)Ge(\lambda)\Big\}\end{aligned} \quad (34b)$$

where v^2, s, and G are given by (25), (26), and (27), respectively.

Equations (34a) and (34b) represent fundamental closed-form expressions for the asymptotic variance and covariance functions, appearing in the joint asymptotically normal distribution of the spectral estimate of an AR process in the presence of additive noise. Their complexity, however, does not allow simple interpretation. Thus, in the next section, we specialize to the case of a first-order AR process and examine the dependence of the asymptotic covariance function acov (λ, ν) on the frequencies λ and ν and on the values of the AR parameter "a."

IV. The AR(1) Case

The first-order AR process has the representation

$$X_n - aX_{n-1} = \epsilon_n \quad (35)$$

where the AR parameter "a" satisfies the condition $-1 < a < 1$. The spectral density for $\phi(\lambda)$ is given by

$$\phi(\lambda) = \frac{\sigma_\epsilon^2/2\pi}{[a^2 - 2a\cos(\lambda) + 1]}. \quad (36)$$

In this case, the asymptotic covariance function acov (λ, ν) can be expressed directly in terms of the parameters a, σ_ϵ^2, σ_W^2. We have

$$\begin{aligned}\operatorname{acov}(\lambda, \nu) = \frac{\phi(\lambda)\phi(\nu)}{\sigma_\epsilon^4}\Big\{&v^2 + 4\pi s\phi(\lambda) \\ &\cdot[\cos(\lambda) - a] + 4\pi s\phi(\nu) \\ &\cdot[\cos(\nu) - a] + (4\pi)^2 G\phi(\lambda)\phi(\nu) \\ &\cdot[\cos(\lambda) - a][\cos(\nu) - a]\Big\} \end{aligned} \quad (37a)$$

$$\begin{aligned}\operatorname{avar}(\lambda) = \frac{\phi^2(\lambda)}{\sigma_\epsilon^4}\Big\{&v^2 + 8\pi s\phi(\lambda)[\cos(\lambda) - a] \\ &+ (4\pi)^2 G\phi^2(\lambda)[\cos(\lambda) - a]^2\Big\} \end{aligned} \quad (37b)$$

where the scalar constants G, s, and v^2 are given by

$$G = \frac{(1 - a^2)}{a^2}\gamma \tag{38}$$

$$s = \frac{\sigma_\epsilon^2}{a^3}\left\{-a^3(1 + a^2) + (1 + a^2)\gamma + 2(1 - a^2)(\sigma_W^2/\sigma_\epsilon^2)\gamma\right\} \tag{39}$$

$$\begin{aligned} v^2 = \sigma_\epsilon^4 &\left\{1 + \frac{3a^2 + 1}{a^2(1 - a^2)} - \frac{2(1 + a^2)^2}{a(1 - a^2)} + \frac{(1 + a^2)^2}{a^4(1 - a^2)}\gamma\right\} + \sigma_\epsilon^2\sigma_W^2\left\{\frac{2a^4 - 1}{a^2} - \frac{4a^2 - a + 3}{a} + \frac{4(1 + a^2)}{a^4}\gamma \right.\\ &\left. + \frac{4}{a^4}(\sigma_W^2/\sigma_\epsilon^2)\gamma - 7\right\} + \sigma_W^4\left\{\frac{(1 + a^2)^2}{a^2}\right\} + 2\sigma_W^6 \end{aligned} \tag{40}$$

where

$$\gamma = \left\{1 + 2(\sigma_W^2/\sigma_\epsilon^2)(1 - a^2) + (\sigma_W^2/\sigma_\epsilon^2)^2(1 - a^4)\right\}.$$

From (37b), we see that the asymptotic variance expression is composed of three terms. The first term $(1/\sigma_\epsilon)^4\phi^2(\lambda)v^2$ represents the contribution due to estimating σ_ϵ^2, the second term

$$(1/\sigma_\epsilon)^4 8\pi s\phi^3(\lambda)[\cos(\lambda) - a]$$

is due to the covariance between the estimates $\hat{\sigma}_\epsilon^2$ and $\hat{a}$, and the third term

$$(1/\sigma_\epsilon)^4(4\pi)^2 G\phi^4(\lambda)[\cos(\lambda) - a]^2$$

represents the contribution due to estimating the AR parameter "a."

We see that even for the AR(1) case, the asymptotic variance expression for the spectral estimate, as given by (37b)-(40), is a complicated function of the process parameters. To provide insight into the relationship between the asymptotic variance and the process parameters a, σ_ϵ^2, σ_W^2, and λ, we evaluated (37b)-(40) for a few parameter cases.

The dependence of the asymptotic variance on the signal-to-noise ratio $\sigma_\epsilon^2/[(1 - a^2)\sigma_W^2]$ is fairly clear from the expressions (37b)-(40) and is expected to be monotonic. Fig. 1 exhibits the normalized asymptotic variance avar$(\lambda)/\phi^2(\lambda)$ as a function of the AR parameter "a" for a wide range of frequencies; the signal-to-noise ratio is set to 1. More interestingly, we wish to compare the normalized asymptotic variance avar$(\lambda)/\phi^2(\lambda)$ to the classical case of AR spectral estimation with no noise. If we set $\sigma_W^2 = 0$ in (37b), we have for our spectral estimate (13), which uses the high-order Yule-Walker equations,

$$\text{avar}(\lambda) = (\phi^2(\lambda)/\sigma_\epsilon^4)\left\{v^2 + 8\pi s\phi(\lambda)[\cos(\lambda) - a] + (4\pi)^2 G\phi^2(\lambda)[\cos(\lambda) - a]^2\right\} \tag{41}$$

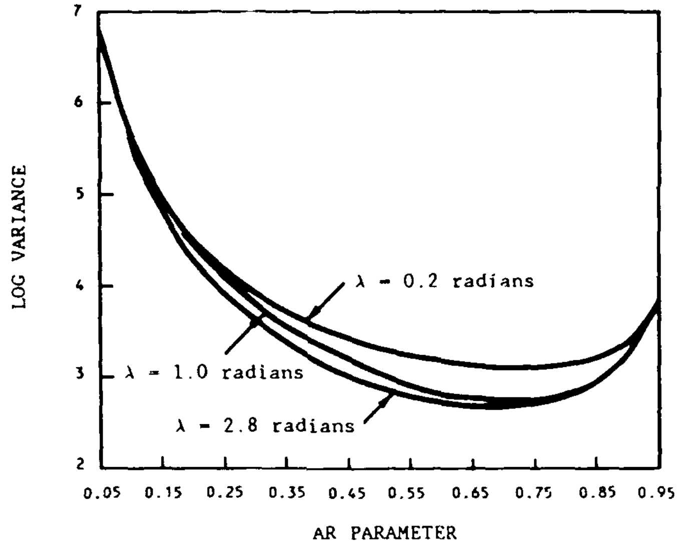

Fig. 1. Normalized asymptotic variance avar$(\lambda)/\phi^2(\lambda)$, (37b), versus AR parameter value; signal-to-noise ratio is one.

where now

$$G = (1 - a^2)/a^2$$

$$s = (\sigma_\epsilon^2/a^3)\left\{(1 + a^2)(1 - a^3)\right\}$$

$$v^2 = \sigma_\epsilon^4 \cdot \left\{1 + \frac{1 + a + 4a^2 + 2a^3 + 6a^4 + 2a^5 + 2a^6}{a^4(1 + a)}\right\}. \tag{42c}$$

On the other hand, for the classical AR spectral estimate, which uses the standard Yule-Walker equations, the asymptotic variance is again given by (41), but with

$$G = (1 - a^2) \tag{43a}$$

$$s = 0 \tag{43b}$$

$$v^2 = 2\sigma_\epsilon^4 \tag{43c}$$

(see Akaike [4]). Fig. 2 plots the ratio of the two asymptotic variances [(41) in conjunction with (42) divided by (41) in conjuction with (43)] as a function of the AR parameter "a" for a wide range of frequencies. It can also be seen from (42) and (43) that this ratio tends to infinity as $a \to 0$ and approaches 5 as $a \to 1$. The larger asymptotic variance of our spectral estimate is basically due to the use of the high-order Yule-Walker equations as seen from the expressions (42) and (43) for G and v^2.

V. Comments

It was noted in the introduction that the results of Lee [8] regarding the asymptotic normality of the AR parameters in a multivariate AR plus noise model are not substantiated. We elaborate on this point in this section. In order to prove the asymptotic normality in question, Lee [8] invokes Theorem 21.1 of Billingsley [9]. This result establishes a central limit theorem for certain nonlinear transformation $\eta_m = f(\cdots, \xi_{m-1}, \xi_m, \xi_{m+1}, \cdots)$ of

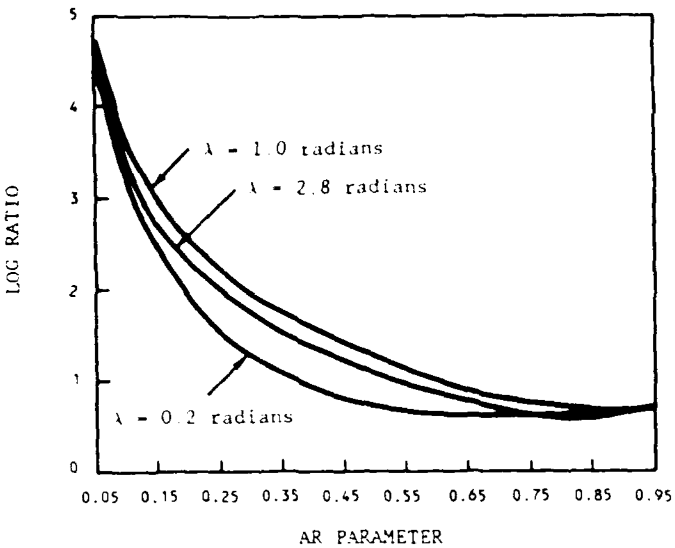

Fig. 2. Ratio of asymptotic variances, (41), with $\sigma_\epsilon^2 = 0$ divided by (41) for the noise-free case versus AR parameter value.

stationary ϕ-mixing processes $\{\xi_i\}$. In order to use it, Lee [8] should have shown that his observation process $\{y_t\}$ (an ARMA(M, M) process) is ϕ mixing and its ϕ-mixing coefficient ϕ_i satisfies

$$\sum_{i=1}^{\infty} \phi_i^{1/2} < \infty$$

(Billingsley [9, Theorem 21.1]). Lee [8, eq. (A10)–(A19)], *unfortunately, is silent on both points. In fact,* Lee could not have proven that these two conditions are satisfied since it is known (see Gastwirth and Rubin [16]) that not even AR processes are ϕ mixing. Consequently, Lee's invocation of Theorem 21.1 of Billingsley [9] is not at all substantiated and neither is his principal result.

VI. Conclusions

The results presented in this paper establish the joint asymptotic normality of AR spectral estimates for the case of noisy observations using the high-order Yule–Walker equations; a precise asymptotic expression for the variance of the limiting distribution is obtained. The paper extends the previous noise-free work of Akaike [4].

The assumption of normality on the processes $\{X_n\}$ and $\{W_n\}$ is not implicitly required to obtain asymptotic normality of the spectral estimate. However, in case these processes are not Gaussian, the asymptotic covariance (17) in Theorem 1 for the estimates $\hat{r}_{N,k}$ will contain an additional term involving the fourth-order cumulant of the process $\{Y_n\}$ (see Brillinger [12]); this term will propagate via Lemmas 2 and 3 to the asymptotic covariance of the spectral estimate $\hat{\phi}_N(\lambda)$.

Appendix A

Proof of Lemma 1: By (9), we have $\boldsymbol{R} = [r_{p+i-j}]_{i,j=1}^{p}$. By Gersch [13], the matrix $\boldsymbol{R}$ is nonsingular; let $\boldsymbol{R}^{-1} = [w_{i,j}]_{i,j=1}^{p}$. Clearly,

$$w_{i,j} = (-1)^{i+j} \det [\boldsymbol{M}_{j,i}]/\det [\boldsymbol{R}]$$

where $\boldsymbol{M}_{j,i}$ is the (i, j)th cofactor of $\boldsymbol{R}$. The matrix $\boldsymbol{R}$ is estimated by $\hat{\boldsymbol{R}}_N = [\hat{r}_{N,p+i-j}]_{i,j=1}^{p}$ where $\hat{r}_{N,k}$ is given by (10). Whenever $\hat{\boldsymbol{R}}_N$ is nonsingular with probability one, we write

$$\hat{\boldsymbol{R}}_N^{-1} = [\hat{w}_{N,i,j}]_{i,j=1}^{p}$$

where

$$\hat{w}_{N,i,j} = (-1)^{i+j} \det [\hat{\boldsymbol{M}}_{N,j,i}]/\det [\hat{\boldsymbol{R}}_N]. \tag{A1}$$

By Parzen [15], $\hat{r}_{N,k}$ converges to r_k almost surely as $N \to \infty$; hence, $\det [\hat{\boldsymbol{R}}_N]$ converges to $\det [\boldsymbol{R}]$ almost surely as $N \to \infty$. Since $\det [\boldsymbol{R}] \neq 0$, as indicated above, we have that with probability one, $\det [\hat{\boldsymbol{R}}_N] \neq 0$ for sufficiently large N. For such an N, (A1) is clearly well defined. Since $\det [\hat{\boldsymbol{M}}_{N,j,i}]$ converges to $\det [\boldsymbol{M}_{j,i}]$ almost surely as $N \to \infty$, we have from (A1) that

$$\hat{w}_{N,i,j} \to (-1)^{i+j} \det [\boldsymbol{M}_{j,i}]/\det [\boldsymbol{R}] = w_{i,j}$$

almost surely as $N \to \infty$.

Since for large N, $\hat{\boldsymbol{a}}_N = \hat{\boldsymbol{R}}_N^{-1} \hat{\boldsymbol{b}}_N$, the almost sure convergence of $\hat{\boldsymbol{a}}_N$ to $\boldsymbol{a}$ as $N \to \infty$ follows from the almost sure convergence of $\hat{\boldsymbol{R}}_N^{-1}$ and $\hat{\boldsymbol{b}}_N$ to $\boldsymbol{R}^{-1}$ and $\boldsymbol{b}$, respectively. □

Proof of Lemma 2: Define the vector $\boldsymbol{v}_N$ by

$$\boldsymbol{v}_N = (\hat{\boldsymbol{a}}_N - \boldsymbol{a}) - \boldsymbol{R}^{-1}\boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}). \tag{A2}$$

By Lemma 1, $\hat{\boldsymbol{R}}_N^{-1}$ exists with probability one for large N. Thus, for large N, we can write

$$\hat{\boldsymbol{a}}_N = \hat{\boldsymbol{R}}_N^{-1} \hat{\boldsymbol{b}}_N. \tag{A3}$$

Using (A3) in (A2), we get

$$\begin{aligned} \boldsymbol{v}_N &= (\hat{\boldsymbol{R}}_N^{-1} \hat{\boldsymbol{b}}_N - \boldsymbol{a}) - \boldsymbol{R}^{-1}\boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}) \\ &= \hat{\boldsymbol{R}}_N^{-1}(\hat{\boldsymbol{b}}_N - \hat{\boldsymbol{R}}_N \boldsymbol{a}) - \boldsymbol{R}^{-1}\boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}). \end{aligned}$$

By the definition of $\boldsymbol{D}$ in the lemma and (15a), it can be seen that $\boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}) = \hat{\boldsymbol{b}}_N - \hat{\boldsymbol{R}}_N \boldsymbol{a}$; thus,

$$\begin{aligned} \boldsymbol{v}_N &= \hat{\boldsymbol{R}}_N^{-1}\boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}) - \boldsymbol{R}^{-1}\boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}) \\ &= (\hat{\boldsymbol{R}}_N^{-1} - \boldsymbol{R}^{-1})\, \boldsymbol{D}(\hat{\boldsymbol{c}}_N - \boldsymbol{c}). \end{aligned}$$

By Lemma 1, $\hat{\boldsymbol{R}}_N^{-1} \to \boldsymbol{R}^{-1}$ almost surely as $N \to \infty$, and by Theorem 1, $N^{1/2}(\hat{\boldsymbol{c}}_N - \boldsymbol{c})$ converges in distribution; thus, by (16a) applied to each component of $\boldsymbol{v}_N$, we have that

$$N^{1/2}|\boldsymbol{v}_N| \to \boldsymbol{0}$$

in probability as $N \to \infty$. The result follows. □

Proof of Lemma 3: By (5), (6), and (12), which exist for large N, we have

$$\begin{aligned} \hat{\sigma}_{N,\epsilon}^2 - \sigma_\epsilon^2 &= -\sum_{j=0}^{p} \hat{a}_{N,j}\hat{r}_{N,j} - (1/\hat{a}_{N,p}) \sum_{j=0}^{p} \hat{a}_{N,j}\hat{r}_{N,p-j} \\ &\quad + \sum_{j=0}^{p} a_j r_j + (1/a_p) \sum_{j=0}^{p} a_j r_{p-j} \\ &= -(-1, \hat{\boldsymbol{a}}_N^T)\, \hat{\boldsymbol{d}}_N - (1/\hat{a}_{N,p})(-1, \hat{\boldsymbol{a}}_N^T)\, \hat{\tilde{\boldsymbol{d}}}_N \\ &\quad + (-1, \boldsymbol{a}^T)\, \boldsymbol{d} + (1/a_p)(-1, \boldsymbol{a}^T)\, \tilde{\boldsymbol{d}}. \end{aligned}$$

Expanding $1/\hat{a}_{N,p}$ in a Taylor series

$$1/\hat{a}_{N,p} = 1/a_p + (1/\gamma_N)^2(\hat{a}_{N,p} - a_p)$$

where $\gamma_N \in (\hat{a}_{N,p}, a_p)$, we obtain

$$\hat{\sigma}^2_{N,\epsilon} - \sigma^2_\epsilon = T_1 + T_2 + T_3$$

with

$$T_1 \equiv -(-1, \hat{a}_N^T)\,\boldsymbol{d}_N + (-1, \boldsymbol{a}^T)\,\boldsymbol{d} \quad \text{(A4a)}$$

$$T_2 \equiv (1/a_p)\left[-(-1, \hat{\boldsymbol{a}}_N^T)\,\hat{\boldsymbol{d}}_N + (-1, \boldsymbol{a}^T)\,\boldsymbol{d}\right] \quad \text{(A4b)}$$

$$T_3 \equiv -(1/\gamma_N)^2(\hat{a}_{N,p} - a_p)(-1, \hat{\boldsymbol{a}}_N^T)\,\hat{\boldsymbol{d}}_N. \quad \text{(A4c)}$$

We first show that

$$N^{1/2}|T_1 + B_1 + B_2| \xrightarrow{P} 0 \quad \text{as } N \to \infty \quad \text{(A5)}$$

where

$$B_1 = ((-1, \hat{\boldsymbol{a}}_N^T) - (-1, \boldsymbol{a}^T))\,\boldsymbol{d} \quad \text{(A6a)}$$

$$B_2 = (-1, \boldsymbol{a}^T)(\boldsymbol{d}_N - \boldsymbol{d}). \quad \text{(A6b)}$$

We have by (A4a) and (A6), after collecting terms,

$$N^{1/2}|T_1 + B_1 + B_2| = N^{1/2}\left|((-1, \boldsymbol{a}^T) - (-1, \hat{\boldsymbol{a}}_N^T))(\boldsymbol{d}_N - \boldsymbol{d})\right|. \quad \text{(A7)}$$

Since by Lemma 1 we have $|(-1, \hat{\boldsymbol{a}}_N^T) - (-1, \boldsymbol{a}^T)| \xrightarrow{P} \boldsymbol{0}$, and by Theorem 1, $N^{1/2}(\boldsymbol{d}_N - \boldsymbol{d})$ converges in distribution as $N \to \infty$, (A5) follows by (16a).

Next we show that

$$N^{1/2}|T_2 + B_3 + B_4| \xrightarrow{P} \boldsymbol{0} \quad \text{as } N \to \infty \quad \text{(A8)}$$

where

$$B_3 = (1/a_p)((-1, \hat{\boldsymbol{a}}_N^T) - (-1, \boldsymbol{a}^T))\,\tilde{\boldsymbol{d}} \quad \text{(A9a)}$$

$$B_4 = (1/a_p)(-1, \boldsymbol{a}^T)(\hat{\boldsymbol{d}}_N - \tilde{\boldsymbol{d}}). \quad \text{(A9b)}$$

We have by (A4b) and (A9), after collecting terms,

$$N^{1/2}|T_2 + B_3 + B_4| = N^{1/2}\left|(1/a_p)\left\{((-1, \boldsymbol{a}^T) - (-1, \hat{\boldsymbol{a}}_N^T))(\hat{\boldsymbol{d}}_N - \boldsymbol{d})\right\}\right|.$$

The result (A8) now follows by the same argument employed earlier for (A7).

Now we show that

$$N^{1/2}|T_3 + B_5| \xrightarrow{P} \boldsymbol{0} \quad \text{as } N \to \infty \quad \text{(A10)}$$

where

$$B_5 = (1/a_p)^2(\hat{a}_{N,p} - a_p)(-1, \boldsymbol{a}^T)\,\tilde{\boldsymbol{d}}.$$

We have

$$N^{1/2}|T_3 + B_5| = \left|\left[(1/a_p)^2(-1, \boldsymbol{a}^T)\,\tilde{\boldsymbol{d}} - (1/\gamma_N)^2(-1, \hat{\boldsymbol{a}}_N^T)\,\hat{\boldsymbol{d}}_N\right] \cdot N^{1/2}(\hat{a}_{N,p} - a_p)\right|.$$

As $N \to \infty$, we have $\hat{\tilde{\boldsymbol{d}}} \to \tilde{\boldsymbol{d}}$ almost surely by Parzen [15], $\hat{\boldsymbol{a}}_N \to \boldsymbol{a}$ almost surely by Lemma 1, and $N^{1/2}(\hat{a}_{N,p} - a_p)$ converges in distribution by Lemma 2. Moreover, $|\gamma_N - a_p| \le |\hat{a}_{N,p} - a_p|$ so that, as $N \to \infty$, $\gamma_N \to a_p$ almost surely since the right side tends to zero by Lemma 1. The result (A10) now follows by (16a).

By (A5), (A8), and (A10), it follows that

$$N^{1/2}\left|(\hat{\sigma}^2_{N,\epsilon} - \sigma^2_\epsilon) - B_1 - B_2 - B_3 - B_4 - B_5\right| \xrightarrow{P} 0 \quad \text{as } N \to \infty. \quad \text{(A11)}$$

By the definition of B_2 and B_4, we have

$$N^{1/2}(B_2 + B_4) = N^{1/2}((-1, \boldsymbol{a}^T)(\boldsymbol{d}_N - \boldsymbol{d}) + (1/a_p)(-1, \boldsymbol{a}^T)(\hat{\boldsymbol{d}}_N - \tilde{\boldsymbol{d}}))$$

and by augmenting the vector $(-1, \boldsymbol{a}^T)$, we write this in the form

$$N^{1/2}(B_2 + B_4) = N^{1/2}((-1, \boldsymbol{a}^T, \boldsymbol{0}^T) + (1/a_p)(\tilde{\boldsymbol{a}}^T, -1, \boldsymbol{0}^T))(\hat{\boldsymbol{c}}_N - \boldsymbol{c})$$
$$= N^{1/2}(\boldsymbol{h}_1^T + \boldsymbol{h}_2^T)(\hat{\boldsymbol{c}}_N - \boldsymbol{c}) \quad \text{(A12)}$$

where

$$\boldsymbol{h}_1 = (-1, \boldsymbol{a}^T, \boldsymbol{0}^T)^T \quad \text{(A13a)}$$

$$\boldsymbol{h}_2 = (1/a_p)(\tilde{\boldsymbol{a}}^T, -1, \boldsymbol{0}^T)^T. \quad \text{(A13b)}$$

Next by the definition of B_1 and B_3, we have

$$N^{1/2}(B_1 + B_3) = N^{1/2}\left\{(\boldsymbol{d} + (1/a_p)\,\tilde{\boldsymbol{d}})^T \cdot ((-1, \hat{\boldsymbol{a}}_N^T)^T - (-1, \boldsymbol{a}^T)^T)\right\}.$$

We need to express this in terms of $(\hat{\boldsymbol{c}}_N - \boldsymbol{c})$. To this end, we note that by Lemma 2, we have

$$N^{1/2}\left|[(-1, \hat{\boldsymbol{a}}_N^T)^T - (-1, \boldsymbol{a}^T)^T] - \begin{bmatrix} \boldsymbol{0}^T \\ \boldsymbol{R}^{-1} \end{bmatrix} D(\hat{\boldsymbol{c}}_N - \boldsymbol{c})\right| \xrightarrow{P} \boldsymbol{0}$$

and thus,

$$N^{1/2}\left|B_1 + B_3 - \boldsymbol{h}_3^T(\hat{\boldsymbol{c}}_N - \boldsymbol{c})\right| \xrightarrow{P} 0 \quad \text{as } N \to \infty \quad \text{(A14)}$$

where

$$\boldsymbol{h}_3 = \left((\boldsymbol{d} + (1/a_p)\,\tilde{\boldsymbol{d}})^T \begin{bmatrix} \boldsymbol{0}^T \\ \boldsymbol{R}^{-1} \end{bmatrix} D\right)^T. \quad \text{(A13c)}$$

Finally, by the definition of B_5, we have

$$N^{1/2}B_5 = N^{1/2}[(1/a_p)^2(\hat{a}_{N,p} - a_p)(-1, \boldsymbol{a}^T)\,\tilde{\boldsymbol{d}}]$$

and by Lemma 2, we have

$$N^{1/2}\left|(\hat{a}_{N,p} - a_p) - [\boldsymbol{R}^{-1}D]_p(\hat{\boldsymbol{c}}_N - \boldsymbol{c})\right| \xrightarrow{P} 0 \quad \text{as } N \to \infty$$

so that

$$N^{1/2}\left|B_5 - \boldsymbol{h}_4^T(\hat{\boldsymbol{c}}_N - \boldsymbol{c})\right| \xrightarrow{P} 0 \quad \text{as } N \to \infty \tag{A15}$$

where

$$\boldsymbol{h}_4 = (1/a_p)^2\left((-1, \boldsymbol{a}^T)\, \bar{d}[\boldsymbol{R}^{-1}\boldsymbol{D}]_p^T\right). \tag{A13d}$$

Thus, with $\boldsymbol{h} = \boldsymbol{h}_1 + \boldsymbol{h}_2 + \boldsymbol{h}_3 + \boldsymbol{h}_4$, the result follows from (A11) using (A12)–(A15).

Appendix B

In this Appendix, the evaluation of the covariance expressions v^2, s, and $\boldsymbol{G}$, stated in (25), (26), and (27), is given. The following lemma is needed.

Lemma B1: We have

a) $$2\pi \int_{-\pi}^{\pi} A^2(e^{-i\lambda})\, e^{im\lambda} \psi^2(\lambda)\, d\lambda$$

$$= \begin{cases} 0; & m > 2p \\ \sigma_W^4 \sum_{k=m-p}^{p} a_k a_{m-k}; & p < m \leq 2p \\ \sigma_W^4 \sum_{k=0}^{p-m} a_k a_{m-k} + 2\sigma_W^2 \sigma_\epsilon^2 a_k \alpha_{k-m} \\ \quad + \sigma_\epsilon^4 \alpha_0^2 \delta_{0,m}; & 0 \leq m \leq p \end{cases}$$

b) $$2\pi \int_{-\pi}^{\pi} A(e^{-i\lambda})\, A(e^{i\lambda})\, e^{im\lambda} \psi^2(\lambda)\, d\lambda$$

$$= \begin{cases} \sigma_\epsilon^2 r_m; & m > p \\ \sigma_\epsilon^2 r_m + \sigma_W^4 \sum_{k=0}^{p-m} a_k a_{k+m} + \sigma_W^2 \sigma_\epsilon^2 \delta_{0,m}; \\ \quad 0 \leq m \leq p \end{cases}$$

c) $$2\pi \int_{-\pi}^{\pi} A(e^{i\lambda})\, A(e^{-i\lambda})\, e^{i(n-m)\lambda} \psi^2(\lambda)\, d\lambda$$

$$= \sigma_W^4 \sum_{k=0}^{p-|n-m|} a_k a_{k+|n-m|} + \sigma_W^2 \sigma_\epsilon^2 \delta_{n,m} + \sigma_\epsilon^2 r_{n-m}$$

where the $\{a_k\}_{k=1}^p$ are the AR parameters of (1), ($a_0 \equiv -1$), the $\{\alpha_k\}_{k=0}^{\infty}$ are the coefficients of the power series expansion for $1/A(z)$, and $\delta_{n,m}$ is the Kronecker delta.

Proof: In the course of the proof, we require the power series expansion

$$1/A(z) = \sum_{k=0}^{\infty} \alpha_k z^k. \tag{B1}$$

Since $A(z)$ is a polynomial of order p with no zeros in $\{z: |z| \leq 1\}$, then the expansion exists and converges uniformly in $|z| \leq 1$.

a) Let

$$T(m) = 2\pi \int_{-\pi}^{\pi} A^2(e^{-i\lambda})\, e^{im\lambda} \psi^2(\lambda)\, d\lambda$$

$$\equiv T_1(m) + T_2(m) + T_3(m) \tag{B2}$$

where by (2) and (3)

$$T_1(m) = (1/2\pi)\, \sigma_W^4 \int_{-\pi}^{\pi} A^2(e^{-i\lambda})\, e^{im\lambda}\, d\lambda \tag{B3}$$

$$T_2(m) = (1/2\pi)\, 2\sigma_W^2 \sigma_\epsilon^2 \int_{-\pi}^{\pi} \frac{A(e^{-i\lambda})}{A(e^{i\lambda})}\, e^{im\lambda}\, d\lambda \tag{B4}$$

$$T_3(m) = (1/2\pi)\, \sigma_\epsilon^4 \int_{-\pi}^{\pi} \frac{e^{im\lambda}}{\left[A(e^{i\lambda})\right]^2}\, d\lambda. \tag{B5}$$

Evaluating each term (B3), (B4), and (B5), we get

$$T_1(m) = (\sigma_W^4/2\pi) \sum_{k=0}^{p} \sum_{j=0}^{p} a_k a_j \int_{-\pi}^{\pi} e^{-i(k+j-m)\lambda}\, d\lambda$$

$$= \begin{cases} 0; & m > 2p \\ \sigma_W^4 \sum_{k=m-p}^{p} a_k a_{m-k}; & p < m \leq 2p \\ \sigma_W^4 \sum_{k=0}^{p} a_k a_{m-k}; & 0 \leq m \leq p. \end{cases} \tag{B6}$$

Next

$$T_2(m) = (2\sigma_W^2 \sigma_\epsilon^2/2\pi) \int_{-\pi}^{\pi} \frac{A(e^{-i\lambda})}{A(e^{i\lambda})}\, e^{im\lambda}\, d\lambda$$

and by (B1) and exchanging of integration and summation, we have

$$T_2(m) = (2\sigma_W^2 \sigma_\epsilon^2/2\pi) \sum_{k=0}^{p} \sum_{j=0}^{\infty} a_k \alpha_j \int_{-\pi}^{\pi} e^{-i(k-j-m)\lambda}\, d\lambda$$

$$= \begin{cases} 0; & m > p \\ 2\sigma_W^2 \sigma_\epsilon^2 a_p \alpha_0; & m = p \\ 2\sigma_W^2 \sigma_\epsilon^2 \sum_{k=m}^{p} a_k \alpha_{k-m}; & 0 \leq m < p. \end{cases} \tag{B7}$$

For $T_3(m)$, we have

$$T_3(m) = (\sigma_\epsilon^4/2\pi) \int_{-\pi}^{\pi} \frac{e^{im\lambda}}{\left[A(e^{i\lambda})\right]^2}\, d\lambda$$

$$= \begin{cases} 0; & m > 0 \\ \sigma_\epsilon^4 \alpha_0; & m = 0 \end{cases} \tag{B8}$$

by (B1). The result follows by (B2)–(B8).

b) Let

$$T(m) = 2\pi \int_{-\pi}^{\pi} A(e^{i\lambda})\, A(e^{-i\lambda})\, e^{im\lambda} \psi^2(\lambda)\, d\lambda$$

$$\equiv T_1(m) + T_2(m) + T_3(m) \tag{B9}$$

where by (2) and (3)

$$T_1(m) = (1/2\pi)\, \sigma_W^4 \int_{-\pi}^{\pi} A(e^{i\lambda})\, A(e^{-i\lambda})\, e^{im\lambda}\, d\lambda \tag{B10}$$

$$T_2(m) = (1/2\pi)\, 2\sigma_W^2 \sigma_\epsilon^2 \int_{-\pi}^{\pi} e^{im\lambda}\, d\lambda \tag{B11}$$

$$T_3(m) = (1/2\pi)\, \sigma_\epsilon^4 \int_{-\pi}^{\pi} \frac{e^{im\lambda}}{A(e^{i\lambda})\, A(e^{-i\lambda})}\, d\lambda. \tag{B12}$$

Proceeding as in part a), we obtain

$$T_1(m) = \begin{cases} 0; & m > p \\ \sigma_W^4 \sum\limits_{k=0}^{p-m} a_k a_{k+m}; & 0 \le m \le p \end{cases} \tag{B13}$$

$$T_2(m) = \begin{cases} 0; & m > 0 \\ 2\sigma_W^2 \sigma_\epsilon^2; & m = 0 \end{cases} \tag{B14}$$

and

$$T_3(m) = \begin{cases} \sigma_\epsilon^2 r_m; & m > 0 \\ \sigma_\epsilon^2 r_0 - \sigma_W^2 \sigma_\epsilon^2; & m = 0 \end{cases} \tag{B15}$$

and the result follows.

c) The proof follows directly from that of part b). □

A. Evaluation of G

For notational convenience in the calculation below, we will not display the dependence on λ for the matrix U. Set

$$B \equiv 2\pi \int_{-\pi}^{\pi} DUD^T \psi^2(\lambda)\, d\lambda; \tag{B16}$$

then $G = R^{-1} B (R^{-1})^T$ by (27). We evaluate explicitly the elements $[B]_{n,m}$ of B. We denote the n,mth element of the matrix DUD^T by

$$[DUD^T]_{n,m} = \sum_{k=0}^{2p} \sum_{j=0}^{2p} d_{n,k} u_{k,j} d_{m,j}$$

$$n, m = 1, 2, \cdots, p$$

where $d_{n,k}$ is the n,kth element of the matrix D and $u_{k,j}$ is the k,jth element of the matrix U. Using the definition of U (15g) and D (Lemma 2), we have

$$\begin{aligned} [DUD^T]_{n,m} &= \sum_{k=n+1}^{n+1+p} a_{p-k+n+1} e^{ik\lambda} \\ &\quad \cdot \sum_{j=m+1}^{m+1+p} a_{p-j+m+1} (e^{ij\lambda} + e^{-ij\lambda}) \\ &= A^2(e^{-i\lambda})\, e^{i(2p+n+m+2)\lambda} \\ &\quad + A(e^{-i\lambda})\, A(e^{i\lambda})\, e^{i(n-m)\lambda}. \end{aligned} \tag{B17}$$

Substituting (B17) into (B16) and using parts a) and c) of Lemma B1, we obtain the final result:

$$\begin{aligned} [B]_{n,m} &= \sigma_W^4 \sum_{k=0}^{p-|n-m|} a_k a_{k+|n-m|} \\ &\quad + \sigma_W^2 \sigma_\epsilon^2 \delta_{n,m} + \sigma_\epsilon^2 r_{n-m}. \end{aligned} \tag{B18}$$

B. Evaluation of s

Put

$$\eta = 2\pi \int_{-\pi}^{\pi} DUh\psi^2(\lambda)\, d\lambda.$$

Then by (26), $s = R^{-1}\eta$ and we evaluate η explicitly. Using the definition of the vector h^T from Lemma 3, we have

$$\begin{aligned} \eta^T &= 2\pi \int_{-\pi}^{\pi} (-1, a^T, 0^T)\, UD^T \psi^2(\lambda)\, d\lambda \\ &\quad - (2\pi/a_p) \int_{-\pi}^{\pi} (a^T, -1, 0^T)\, UD^T \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} (d^T + (1/a_p)\, d^T) \begin{bmatrix} 0^T \\ R^{-1} \end{bmatrix} \\ &\quad \cdot DUD^T \psi^2(\lambda)\, d\lambda \\ &\quad + (2\pi/a_p^2) \int_{-\pi}^{\pi} ((-1, a^T)\, d)\, [R^{-1} D]_p \\ &\quad \cdot UD^T \psi^2(\lambda)\, d\lambda \\ &\equiv \beta_1^T + \beta_2^T + \beta_3^T + \beta_4^T. \end{aligned} \tag{B19}$$

For β_1^T using the definition of U and D, the nth element of $(-1, \tilde{a}^T, 0^T)\, UD^T$ is given by

$$\begin{aligned} ((-1, \tilde{a}^T, 0^T)\, UD^T)_n &= A(e^{i\lambda})\, A(e^{-i\lambda})\, e^{i(p-n+1)\lambda} \\ &\quad + A^2(e^{i\lambda})\, e^{i(p+n+1)\lambda}. \end{aligned}$$

Using parts a) and b) of Lemma B1, we then have

$$(\beta_1^T)_n = -\sigma_\epsilon^2 r_{p+n+1} - \sigma_W^4 \sum_{k=n+1}^{p} a_k a_{p+n+1-k}. \tag{B20}$$

For β_2^T using the definition of U and D, the nth element of $(a^T, -1, 0^T)\, UD^T$ is given by

$$\begin{aligned} ((a^T, -1, 0^T)\, UD)_n &= A^2(e^{-i\lambda})\, e^{i(2p+n+1)\lambda} \\ &\quad + A(e^{-i\lambda})\, A(e^{i\lambda})\, e^{-i(n+1)\lambda}. \end{aligned}$$

Using parts a) and b) of Lemma B1, we obtain

$$(\beta_2^T)_n = -(\sigma_W^4/a_p) \sum_{k=n+1}^{p} a_k a_{k-n+1} - (\sigma_\epsilon^2/a_p)\, r_{n+1}. \tag{B21}$$

For β_3^T using (B16), it follows that

$$\beta_3^T = (d^T + (1/a_p)\, d^T) \begin{bmatrix} 0^T \\ R^{-1} \end{bmatrix} B \tag{B22}$$

where B is given in (B18). For β_4, we can write

$$\begin{aligned} \beta_4^T &= (2\pi/a_p^2) \int_{-\pi}^{\pi} ((-1, a^T)\, d) \\ &\quad \cdot [R^{-1}]_p DUD^T \psi^2(\lambda)\, d\lambda \end{aligned}$$

and by (B16), it follows that

$$\boldsymbol{\beta}_4^T = (1/a_p)^2 \left[\sum_{j=0}^{p} a_j r_{p-j} \right] [\boldsymbol{R}^{-1}]_p \boldsymbol{B}. \tag{B23}$$

Thus, $s = \boldsymbol{R}^{-1}\boldsymbol{\eta}$ where $\boldsymbol{\eta}$ is given by (B19)–(B23).

C. Evaluation of v^2

By (25) and the definition of $\boldsymbol{h} = \boldsymbol{h}_1 + \boldsymbol{h}_2 + \boldsymbol{h}_3 + \boldsymbol{h}_4$ [cf. (A13)], we write

$$\begin{aligned} v^2 &= 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}^T \boldsymbol{U} \boldsymbol{h}_1 \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}^T \boldsymbol{U} \boldsymbol{h}_2 \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}^T \boldsymbol{U} \boldsymbol{h}_3 \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}^T \boldsymbol{U} \boldsymbol{h}_4 \psi^2(\lambda)\, d\lambda \\ &\equiv T_1 + T_2 + T_3 + T_4. \end{aligned} \tag{B24}$$

For T_1, we have

$$\begin{aligned} T_1 &= 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}_1^T \boldsymbol{U} \boldsymbol{h}_1 \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}_2^T \boldsymbol{U} \boldsymbol{h}_1 \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}_3^T \boldsymbol{U} \boldsymbol{h}_1 \psi^2(\lambda)\, d\lambda \\ &\quad + 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}_4^T \boldsymbol{U} \boldsymbol{h}_1 \psi^2(\lambda)\, d\lambda \\ &\equiv T_{11} + T_{12} + T_{13} + T_{14}. \end{aligned} \tag{B25}$$

Using (A13a)–(A13d) for $\boldsymbol{h}_1$, $\boldsymbol{h}_2$, $\boldsymbol{h}_3$, and $\boldsymbol{h}_4$, applying Lemma B1, and using $\boldsymbol{\beta}_1$ from (B19), we obtain

$$\begin{aligned} T_{11} &= \sigma_W^4 + \sigma_\epsilon^4 + \sigma_\epsilon^2 r_0 + \sigma_W^4 \sum_{k=0}^{p} a_k^2 + 2\sigma_W^2 \sigma_\epsilon^2 \\ &\quad \cdot \sum_{k=0}^{p} a_k \alpha_k - \sigma_W^2 \sigma_\epsilon^2 \end{aligned} \tag{B26a}$$

$$\begin{aligned} T_{12} &= \Big\{ (1/a_p) \sigma_\epsilon^2 r_p - \sigma_W^4 a_p \\ &\quad + \sigma_W^4 \sum_{k=0}^{p} a_k a_{p-k} - 2\sigma_W^2 \sigma_\epsilon^2 a_p \Big\} \end{aligned} \tag{B26b}$$

$$T_{13} = \left(\boldsymbol{d}^T + (1/a_p)\, \tilde{\boldsymbol{d}}^T\right) \begin{bmatrix} \boldsymbol{0}^T \\ \boldsymbol{R}^{-1} \end{bmatrix} \boldsymbol{\beta}_1 \tag{B26c}$$

$$T_{14} = (1/a_p)^2 \left[\sum_{j=0}^{p} a_j r_{p-j} \right] [\boldsymbol{R}^{-1}]_p \boldsymbol{\beta}_1. \tag{B26d}$$

Similarly, for T_2, we write

$$T_2 = T_{21} + T_{22} + T_{23} + T_{24}. \tag{B27}$$

Using (A13a)–(A13d), Lemma B1, and $\boldsymbol{\beta}_2$ from (B19), we get

$$T_{21} = T_{12} \tag{B28a}$$

$$T_{22} = (1/a_p^2) \left\{ \sigma_W^4 a_p^2 + \sigma_W^4 \sum_{k=0}^{p} a_k^2 + \sigma_\epsilon^2 r_0 - \sigma_\epsilon^2 \sigma_W^2 \right\} \tag{B28b}$$

$$T_{23} = \left(\boldsymbol{d}^T + (1/a_p)\, \tilde{\boldsymbol{d}}^T\right) \begin{bmatrix} \boldsymbol{0}^T \\ \boldsymbol{R}^{-1} \end{bmatrix} \boldsymbol{\beta}_2 \tag{B28c}$$

$$T_{24} = (1/a_p)^3 \left[\sum_{j=0}^{p} a_j r_{p-j} \right] [\boldsymbol{R}^{-1}]_p \boldsymbol{\beta}_2. \tag{B28d}$$

Using (A13c) for $\boldsymbol{h}_3$, we write

$$T_3 = 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}^T \boldsymbol{U} \boldsymbol{D}^T \begin{bmatrix} \boldsymbol{0}^T \\ \boldsymbol{R}^{-1} \end{bmatrix} \left(\boldsymbol{d} + (1/a_p)\, \tilde{\boldsymbol{d}}\right) \psi^2(\lambda)\, d\lambda,$$

but, by (26), we have that

$$\boldsymbol{s}^T = 2\pi \int_{-\pi}^{\pi} \boldsymbol{h} \boldsymbol{U} \boldsymbol{D}^T (\boldsymbol{R}^{-1})^T \psi^2(\lambda)\, d\lambda;$$

thus,

$$T_3 = (0, \boldsymbol{s}^T)\left(\boldsymbol{d} + (1/a_p)\, \tilde{\boldsymbol{d}}\right). \tag{B29}$$

Using (A13d) for $\boldsymbol{h}_4$, we write

$$\begin{aligned} T_4 &= 2\pi \int_{-\pi}^{\pi} \boldsymbol{h}^T \boldsymbol{U} [\boldsymbol{D}^T (\boldsymbol{R}^{-1})^T]_p (1/a_p)^2 \\ &\quad \cdot \left[\sum_{j=0}^{p} a_j r_{p-j} \right] \psi^2(\lambda)\, d\lambda. \end{aligned}$$

As above, by (26), we obtain

$$T_4 = (\boldsymbol{s}^T)_p (1/a_p)^2 \left[\sum_{j=0}^{p} a_j r_{p-j} \right]. \tag{B30}$$

Thus, $v^2 = \Sigma_{i=1}^4 T_i$ where the expressions for T_i are given by (B25)–(B30).

References

[1] M. Kaveh and G. R. Cooper, "An empirical investigation of the properties of the autoregressive spectral estimator," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 313-323, 1976.

[2] N. Beamish and M. B. Priestly, "A study of autoregressive and window spectral estimation," *Appl. Statist.*, vol. 30, pp. 41-58, 1981.

[3] H. B. Mann and A. Wald, "On the statistical treatment of linear stochastic difference equations," *Econometrica*, vol. 11, pp. 173-220, 1943.

[4] E. H. Akaike, "Power spectrum estimation through autoregressive model fitting," *Ann. Inst. Statist. Math.*, vol. 21, pp. 407-419, 1969.

[5] K. N. Berk, "Consistent autoregressive spectral estimates," *Ann. Statist.*, vol. 2, pp. 489-502, 1974.

[6] A. M. Walker, "Some consequences of superimposed error in time series analysis," *Biometrika*, vol. 47, pp. 33-43, 1960.

[7] M. Pagano, "Estimation of models of autoregressive signal plus white noise," *Ann. Statist.*, vol. 2, pp. 99-108, 1974.

[8] T. S. Lee, "Large sample identification and spectral estimation of noisy multivariate autoregressive processes," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 76-82, 1983.
[9] P. Billingsley, *Convergence of Probability Measures.* New York: Wiley, 1968.
[10] D. F. Gingras, "Asymptotic properties of high-order Yule-Walker estimates of the AR parameters of an ARMA time series," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 1095-1101, 1985.
[11] C. R. Rao, *Linear Statistical Inference and Its Applications.* New York: Wiley, 1965.
[12] D. R. Brillinger, *Time Series, Data Analysis and Theory.* New York: Holt, Rinehart and Winston, 1975.
[13] W. Gersch, "Estimation of the autoregressive parameters of a mixed autoregressive moving average time series," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 583-588, 1970.
[14] R. J. Serfling, *Approximation Theorems of Mathematical Statistics.* New York: Wiley, 1980.
[15] E. Parzen, "An approach to time series analysis," *Ann. Math. Statist.*, vol. 32, pp. 951-989, 1961.
[16] J. L. Gastwirth and H. Rubin, "The asymptotic distribution theory of the empiric CDF for mixing stochastic processes," *Ann. Statist.*, vol. 3, pp. 809-824, 1975.

Donald F. Gingras (M'77) received the B.S. degree from San Diego State University, San Diego, CA, in 1968, the M.S. degree from Northeastern University, Boston, MA, in 1971, and the Ph.D. degree from the University of California, San Diego, in 1986, all in electrical engineering.

From 1968 to 1971 he was with Raytheon, Submarine Signal Division, Portsmouth, RI. Since 1971 he has been with the Naval Ocean Systems Center, San Diego, CA. Since 1986 he has also been affiliated with the University of California, San Diego, as a Lecturer in Electrical Engineering. His research interests include spatial array processing, time series analysis, and statistical signal processing.

Elias Masry (S'64-M'68-SM'83-F'86) [redacted] He received the B.Sc. and M.Sc. degrees from the Technion-Israel Institute of Technology, Haifa, in 1963 and 1965, respectively, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, in 1966 and 1968, respectively, in electrical engineering.

From 1955 to 1959 he served in the Israel Defence Forces. From 1963 to 1965 he was a Teaching Assistant at the Technion. Since 1968 he has been on the Faculty of the University of California at San Diego, La Jolla, and is currently a Professor of Electrical Engineering. His research interests include time series analysis, sampling design, and adaptive filtering.

Dr. Masry was the Associate Editor for Stochastic Processes of the IEEE TRANSACTIONS ON INFORMATION THEORY from 1980 to 1983.